DOCUMENT RESUME

ED 262 101 TM 850 589

AUTHOR Smylie, Mark
TITLE Improving the Accuracy of Teacher Self-Evaluation through Staff Development.
PUB DATE Mar 85
NOTE 29p.; Paper presented at the Annual Meeting of the American Educational Research Association (69th, Chicago, IL, March 31-April 4, 1985).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Classroom Observation Techniques; Classroom Research; Elementary Secondary Education; Feedback; *Inservice Teacher Education; *Lesson Observation Criteria; Program Evaluation; *Self Evaluation (Individuals); Student Reaction; *Teacher Evaluation; Teacher Improvement; Teacher Workshops
IDENTIFIERS *Effective Use of Time Program; *Stallings Observation Instrument

ABSTRACT

This study examined impact of a staff development program on increasing the accuracy of teachers' self-evaluation of classroom performance. Teachers were provided with specific formal feedback through a program called The Effective Use of Time Program (EUOT). The program had four components: (1) observation in the classroom and feedback about performance; (2) information from research applied in the context of a teacher's specific situation; (3) opportunities for guided practice; and (4) observation and feedback following workshops to assess improvement. To test EUOT, data from 40 elementary, junior high, and secondary teachers were examined. These observations, using the Stallings Observation Instrument (SOI), provided a message of actual teacher performance. Four variables were selected for comparision from self-rating and observation data: (1) managerial and organization efficiency; (2) student discipline; (3) interactive instruction; and (4) monitoring. Results showed that teachers improved their actual performance and gave themselves higher ratings at the end of the program. In general, teachers were no more accurate in assessing their performance after training, although junior high teachers showed significant relationships between posttraining self-ratings and observed behavior. Several explanations are given for failure of data to show wider improvements in accuracy of self-assessment, and five tables are included to show results. (LMO)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED262101

IMPROVING THE ACCURACY OF TEACHER SELF-EVALUATION THROUGH STAFF DEVELOPMENT

Mark Smylie

Peabody College
Vanderbilt University

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✔ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Smylie, M.

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Illinois, April, 1985

TM 850 589

IMPROVING THE ACCURACY OF TEACHER SELF-EVALUATION THROUGH STAFF DEVELOPMENT

In order to improve their instructional effectiveness, teachers need information about their classroom performance. Such information comes from a variety of sources. These sources include formal evaluations, usually conducted at the building level by the principal, informal feedback from teaching colleagues, and the teacher's own experience with students and self-perceptions of the effectiveness of their performance.

The problem is that these sources generally do not provide much relevant and useful information to help teachers assess and improve their teaching performance. The usual approach to formal evaluation is for the buillding principal to complete a checklist type form indicating impressions of certain teacher characteristics or mastery of certain skills that may have little relationship to promoting student learning (Weisenstein 1976; Hawley & Rosenholtz 1984). These evaluations occur irregularly and infrequently (Dreeben 1970; Dornbusch & Scott 1975; Natriello & Dornbusch 1980-1981) and leave teachers with little sense of the effectiveness of their performance or of the steps they need to take for improvement (Lortie 1975; Natriello & Dornbusch 1980-1981; Natriello 1984).(1)

Informal collegial evaluation and feedback may be one of the most potentially productive sources of information to improve teacher practice (Little 1982; Fullan 1982; Rosenholtz & Smylie 1984). However, collegial feedback about performance is rare in most schools and is restricted by the physical isolation of teachers and norms of autonomy (Griffin 1983). For example, Lortie (1975) found in his

sample of elementary and secondary school teachers that 45% reported no contact with other teachers during the course of the workday and only 32% reported having occasional contact (see also Tye & Tye 1984). Glidewell and his colleagues (1983) identified autonomy as something of a moral imperative among teachers in their study. These teachers felt constrained against offering suggestions about even the most routine matters to other teachers in their schools.

In the absence of relevant and useful feedback from formal and collegial evaluation, most teachers are left to rely on self-evaluation of their classroom experiences (see Buchmann & Schwille 1983; Rosenholtz & Smylie 1984). As measured by self-reports, teachers' perceptions generally do not provide accurate assessments of their actual performance (Hook & Rosenshine 1979; Jones 1982; Wheeler & Knoop 1982). Without being able to accurately assess their own performance, indeed, without constructive feedback from external sources, teachers are placed in a position of reacting by trial and error to situations that do not go as planned or to classroom crises.

The purpose of this study is to examine the impact of a staff development program on increasing the accuracy of teachers' self-evaluations of their classroom performance. The program, The Effective Use of Time Program, provides teachers with specific formal feedback about their teaching and presents a variety of opportunities for teachers to experiment with new classroom strategies and to informally assess and discuss their experiences in an interactive collegial environment. It is hypothesized that this program will help teachers become more aware of their behavior and increase the accuracy of their self-assessments of classroom performance.

Description of the Staff Development Program

The Effective Use of Time Program (EUOT) is based on a four-component accountability model for staff development (Stallings, 1985). In the first component, teachers are observed in their classrooms and provided detailed objective feedback about their performance. Areas of deficiency are identified and teachers are assisted to set their own individual goals for improvement. The second component is designed to provide teachers with information from research findings on effective teaching. These findings are presented in the context of broader theory of instructional practice and linked to the individual teacher's specific classroom situation. The third component allows opportunities for guided practice. Here, teachers take research findings and different instructional strategies provided by other teachers and, under the guidance of the workshop leader and with the assistance of teachers in the workshop group, work to adapt and experiment with them according to the teacher's own teaching style and classroom context. Within the group, teachers assess and discuss their experiences with implementing new strategies. Further, each teacher is provided additional feedback about their performance from peer observations. In the fourth component, after the workshops conclude, teachers are again observed in their classrooms and provided the same types of objective feedback presented at the beginning of the program to assess their levels of improvement.

The program consists of six workshops. The first workshop is devoted to presentation of general research findings and discussion of pretraining observational data. In the sixth workshop, teachers are provided posttraining feedback about the changes they made during the course of the program. The other four workshops focus on strategies

to improve classroom management, student discipline, lesson development and implementation, and interactive instruction and monitoring independent student work. The goal is to provide teachers with methods that will help them immediately in the classroom to use their time and their students' time in academically-productive ways.

Each workshop group is headed by a certified trainer who serves as a facilitator of discussion and experimentation among teacher members. The teachers themselves are encouraged to serve as resources and advisors for one another. To encourage discussion and to build a sense of trust and confidence among teachers in the group, membership is limited to six to eight teachers. To make that discussion more relevant to participants, workshop groups are structured to include teachers of similar grade levels (e.g. elementary, junior high/middle school, secondary) and subject areas (e.g. academic as opposed to vocational education or fine arts). Finally, to provide teachers time to experiment with new strategies and reflect upon their experiences, workshop sessions are scheduled one to two weeks apart.

It seems likely that various aspects of this program could serve to increase teachers' abilities to accurately evaluate their own performance. First, teachers are provided with detailed objective data about their actual classroom performance. And, they are able to compare their performance to research findings that suggest ranges of effective levels of behavior (for a description of these levels, see Stallings 1985; Simons 1985). Second, teachers are encouraged to examine their own practice in light of their specific classroom situations to determine how different strategies might or might not apply. These assessments are made in the interactive context of the work group. Third, workshop activity pivots on discussion of

strategies teachers currently use, experimentation with new ideas, and assessments of teachers' experiences putting those new ideas into practice. Participants commit to try new ideas, return to the classroom to implement them, and then report the results back to the group. Each teacher with other group members discuss successes and failures and try to determine why certain outcomes may have resulted. Finally, the opportunity to conduct peer observations provides additional feedback to the teacher being observed about classroom performance. Further, and of equal importance, these observations put teachers in a position to reflect upon their practice as they observe their colleagues.

Methodology

To test whether the EUOT Program serves to increase the accuracy of teachers' self-evaluation of their classroom performance, data were examined from 40 elementary, junior high/middle school, and secondary teachers who participated in the program between 1982 and 1984.(2) Each teacher completed a performance self-rating form during the first workshop before they received feedback from their classroom observations. During the sixth and final workshop, these teachers completed the same self-rating form before they received feedback from their posttraining observations. The form asks teachers to assess their performance in various one-item categories on a scale of 1 to 4 where 1 = Needs Improvement and 4 = Very Effective.

Classroom observation data provide a measure of actual teacher performance. These data were collected during three classroom observations before training and three classroom observations after training with the Stallings Observation Instrument (SOI). The SOI provides data for both teachers and students on proportions of time

spent in and proportions of interactions related to classroom organization and management, student discipline, interactive instruction and monitoring (see Stallings & Geisen 1977 for a detailed description of this instrument, its validity and its reliability). Proportions of time spent in various activities were derived from five short scans of the classroom equally spaced throughout the class period. In all, over the course of these observations, 15 samples of activity were recorded. Classroom interactions were coded at virtually the speed of speech during five timed 5-minute periods, also equally spaced across the class period. During the three classroom observations between 750 and 900 interactions for each teacher were recorded. For the purpose of this study, we rely as much as possible on interaction data as opposed to the activity data collected from scans. We do so because the volume of data points about interactions are greater and provide a more stable data base for analysis.

Four variables were selected for comparison from both self-rating and observation data. Teachers' self-ratings on managerial and organization efficiency were matched with the proportion of all observed statements that dealt with management and organization. Their self-ratings on student discipline were matched with the proportion of all observed statements related to student behavior. Self-ratings on interactive instruction were matched with the proportion of observed statements that were academic. Finally, teacher self-ratings on monitoring were matched with the difference between the observed proportions of time students spent doing seatwork and the teacher spent monitoring that seatwork. We chose to represent the observed monitoring variable this way to account for the differing amounts of seatwork teachers assigned to their students. Interaction

data would not provide this distinction.

The observational variables had to be adjusted to reflect levels of relative effectiveness so that they might be compared to the teachers' self-ratings of effective performance. To make this computational adjustment, individual scores on each of the first three observational variables were subtracted from criterion scores established in previous research (Stallings, Cory, Fairweather, & Needels 1978). These criterion scores represent levels of performance significantly related to student achievement (Stallings et al. 1978; Stallings 1980). For classroom organization and management, the criterion is 15% or less of all observed statements. For student discipline, the criterion is 3% or less of all statements. For interactive instruction, the criterion is 80% or more of all statements.

Thus, in order to show relative levels of effective performance, teachers' observation scores for classroom management were subtracted from 15%. Teachers' observation scores for student discipline were subtracted from 3%. Finally, teachers' scores for instruction were subtracted from 80% and then converted to absolute values to indicate proper direction with respect to the criterion level. In this way, low adjusted scores represent low levels of effective performance and high adjusted scores represent high levels of effective performance (Thieme-Busch & Prom 1983).

Levels of effective monitoring were computed by subtracting the proportion of time students spent doing seatwork from the proportion of time teachers spent monitoring that seatwork. The working assumption here is that academically effective teachers do not attend to other duties while their students work independently. Instead,

they actively monitor that activity to help keep students on-task and correct any errors students may be making (Stallings 1980). Thus, higher scores on this adjusted measure indicate that teachers are involved more in monitoring their students' seatwork than teachers with lower scores.

Means and standard deviations were calculated and one-tail T-tests were computed for each of the four variables from self-ratings and adjusted observation scores both before and after training to indicate directions and significance of change. Then, analyses of variance were conducted to test whether teachers were more accurate in assessing their own behavior after training than before (Nie, Hull, Jenkins, Steinbrenner & Bent 1975). It was anticipated that the relationship between the variance in posttraining self-ratings and posttraining observed classroom performance would be more significant than the relationship between the variance of pretraining self-ratings and observed performance.

Analyses were conducted for classroom organization and management, student discipline, interactive instruction, and monitoring to determine if teachers' were more accurate in assessing their own performance in some areas but not others. And, analyses were conducted on subsamples of the 40 teachers to test whether elementary, junior high/middle school, or secondary teachers might be better self-evaluators.

Results

Pretraining and posttraining means for self-ratings and adjusted observation scores are presented in Table 1. Changes in these means suggest that overall, teachers improved their actual performance and gave themselves higher ratings at the end of the program.(3) However,

few of these changes were statistically significant.

Analyses of variance indicate that in general the teachers were no more accurate in assessing the effectiveness of their classroom performance after training than before. Results shown in Tables 2, 3, 4, and 5 reveal insignificant relationships between self-ratings and observed behavior both before and after training in all four areas of performance. Insignificant relationships were found for teachers in the total sample and for subsamples of elementary, junior high/middle school, and high school teachers.

There are several exceptions. Significant relationships were found between posttraining self-ratings and observed behavior for junior high/middle school teachers in the areas of classroom management ($p=.026$) and interactive instruction ($p=.044$) (see Tables 2 and 4). Significant relationships were also found between posttraining self-ratings and observed performance in the area of interactive instruction ($p=.007$) and between pretraining and posttraining self-ratings and observed performance in the area of monitoring ($p=.001$; $p=.033$) for the total sample (see Tables 4 and 5).

Discussion

Teachers assess the effectiveness of their performance in large part by the performance of their students. Evidence exists that influence in classroom interaction is bidirectional (Fiedler 1975; Brophy & Good 1974). The teacher acts; students respond; the teacher uses student responses to evaluate the effectiveness of the action and determine how next to proceed (see also Green 1983).

This staff development program requires that teachers experiment with new ideas and strategies in the classroom and study their experiences. Such experimentation introduces change in routine for

both teachers and students. How students respond to the change that is introduced will indicate to the teacher the effectiveness of that change. The stronger the student response, either positive or negative, the stronger the message to the teacher about his or her performance.

Indeed, it is likely that teachers' ability to accurately assess their own performance is related to the magnitude of student reaction. One would anticipate that teachers of more reactive students would be more aware of their own behavior. And, there is some evidence to suggest that secondary students, and particularly junior high students, are more reactive than elementary students (see e.g., Fiedler 1975). This phenomenon may help explain why significant relationships between posttraining self-ratings and observations were found for junior high/middle schol teachers but not for elementary and high school teachers.

Failure of the data to show wider improvements in the accuracy of teachers' self-assessments may be due to a variety of other factors as well. First, the self-rating instrument may not be sensitive enough to assess finer gradations in teachers' perceptions of the effectiveness of their performance. Second, the specific feedback, discussion of individual practice, and presentation of research findings about effective teaching may alter teachers' frames of reference, raise their standards for effective performance, and make them more self-critical of their own behavior.

Third, it is possible that over the course of the workshops, as teachers discuss their individual situations, they indeed become more accurate in assessing their own performance. But those assessments may be more precise in relation to the specific contexts of their

individual classrooms instead of in relation to broader sets of performance criteria used in the program and for this analysis. This raises important questions about the applicability of criteria developed from studies of teachers in one context (Stallings et al. 1978) to other populations of teachers in different contexts.

Finally, even though the staff development program provides a great deal of specific feedback about individual practice and performance, it does not offer a systematic set of strategies for teachers to use to assess their own performance. The presence of such strategies seems to be key in those few studies that do find significant relationships between self-assessments and measures of actual performance (see Irvine 1983; Newfield 1980).

Notes

1. More productive, interactive evaluation processes have been identified in studies of effective schools and of effective evaluation and staff development programs (see e.g., Wise, Darling-Hammond, McLaughlin, & Bernstein 1984; Azumi & Madhere 1983; Wynne 1980; Wellisch, MacQueen, Carriere, & Duck 1978; Brookover & Lezotte 1979; Good & Grouws 1979; Stallings 1980). While this research adds to the body of evidence about more effective ways to help improve practice, these processes seem to be exceptions rather than the rule in most school districts.

2. This sample includes all teachers in eight workshop groups held in California and Tennessee for whom complete data are available.

3. It should be noted that the large differences between the pretraining and posttraining adjusted monitoring scores for elementary teachers and the pretraining and posttraining monitoring scores for the other subsamples of teachers is probably due to greater amounts of small group work in elementary classrooms. Elementary teachers are more likely to assign independent seatwork to one group and then move to another group for instruction than junior high/middle school or high school teachers (Stallings 1980). For this reason, the amount of time students spend doing seatwork may be greater than the time their teachers spend monitoring that work.

References

Azumi, J.E. & Madhere, S. (1983, April). Professionalism, power and performance: The relationships between administrative control, teacher conformity and student achievement. Paper presented at the annual meeting of the American Educational Research Association, Montreal.

Brookover, W.B. & Lezotte, L.W. (1979). Changes in school characteristics coincident with changes in student achievement. E. Lansing: Michigan State University, Institute for Research on Teaching.

Brophy, J.E. & Good, T.L. (1974). Teacher-student relationships: Causes and consequences. New York: Holt, Rinehart & Winston.

Buchmann, M. & Schwille, J. (1983). Education: The overcoming of experience. American Journal of Education, 30-51.

Dornbusch, S.M. & Scott, W.R. (1975). Evaluation and the exercise of authority. San Francisco, CA: Jossey-Bass.

Dreeben, R. (1970). The nature of teaching. Glenview, IL: Scott Foresman.

Fiedler, M.L. (1975). Bidirectionality of influence in classroom interaction. Journal of Educational Psychology, 67, 735-744.

Fullan, M. (1982). The meaning of educational change. New York: Teachers College Press.

Glidewell, J.C.; Tucker, S., Todt, M.; & Cox, S. (1983). Professional support systems: The teaching profession. In A. Nadler, J.D. Fisher, & B.M. DePaulo (Eds.), Applied research in help-seeking and reactions to aid. New York: Academic Press.

Good, T.L. & Grouws, D. (1979). The Missouri mathematics effectiveness project. Journal of Educational Psychology, 71, 355-382.

Green, J.L. (1983). Research on teaching as a linguistic process: A state of the art. In E.W. Gordon (Ed.), Review of research in education (Vol. 10). Washington, DC: American Educational Research Association.

Griffin, G.A. (1983). Implications of research for staff development programs. Elementary School Journal , 83, 414-425.

Hawley, W.D. & Rosenholtz, S.J. (1984). Good schools: What research says about improving student achievement. Peabody Journal of Education, 61 (4), 1-178.

Hook, C.M. & Rosenshine, B. (1979). Accuracy of teacher reports of their classroom behavior. Review of Educational Research, 49, 1-11.

Irvine, J.J. (1983). The accuracy of pre-service teachers' assessments of their classroom behaviors. Journal of Research and Development in

Education, 17, 25-31.

Jones, T.W. (1982). The validity of self-evaluation of competence by special education teacher trainees. Texas Tech Journal of Education, 9, 23-27.

Little, J.W. (1982). Norms of collegiality and experimentation: Workplace conditions of school success. American Educational Research Journal, 19, 325-340.

Lortie, D.C. (1975). Schoolteacher: A sociological study. Chicago: University of Chicago Press.

Natriello, G. (1984). Teachers' perceptions of the frequency of evaluation and assessments of their effort and effectiveness. American Educational Research Journal, 21, 579-595.

Natriello, G. & Dornbusch, S.M. (1980-1981). Pitfalls in the evaluation of teachers by principals. Administrator's Notebook, 29 (6).

Newfield, J. (1980). Accuracy of teacher reports: Reports and observations of specific classroom behaviors. Journal of Educational Research, 74, 78-82.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K., & Bent, D.H. (1975). Statistical package for the social sciences (2nd edition). New York: McGraw-Hill.

Rosenholtz, S.J. & Smylie, M.A. (1984). Teacher compensation and career ladders. Elementary School Journal, 85, 149-166.

Simons, S. (1985, April). Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Stallings, J. (1980). Allocated academic learning time revisited, or beyond time on task. Educational Researcher, 9 (11), 11-16.

Stallings, J. (April, 1985). How effective is an analytic approach to staff development on teacher and student behavior? Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Stallings, J.; Cory, R.; Fairweather, J.; & Needels, M. (1978). A study of basic reading skills taught in secondary schools. Menlo Park, CA: SRI.

Stallings, J. & Giesen, P. (1977, February). The study of reliability in observational data (Occasional Paper No. 19). Bloomington, IN: Phi Delta Kappa.

Thieme-Busch, C.A. & Prom, S.E. (1983, April). Impact of teacher use of time training on student achievement. Paper presented at the annual meeting of the American Educational Research Association, Montreal.

Tye, K.A. & Tye, B.B. (1984). Teacher isolation and school reform. Phi

Delta Kappan , 65, 319-322.

Weisenstein, G.R. (1976). Teacher evaluation: The principal's role. OSCC Bulletin, 20, 3-23.

Wellisch, J.B., MacQueen, A.H., Carriere, R.A., & Duck, G.A. (1978). School management and organization in successful schools. Sociology of Education, 51, 211-226.

Wheeler, A.W. & Knoop, H.R. (1982). Self, teacher, and faculty assessments of student teaching performance. Journal of Educational Research, 75, 178-181.

Wise, A.E., Darling-Hammond, L., McLaughlin, M.W., & Bernstein, H.T. (1984). Teacher evaluation: A study of effective practices. Santa Monica, CA: Rand.

Wynne, E.A. (1980). Looking at schools: Good, bad, and indifferent, Lexington, MA: D.C. Heath.

TABLE 1. Means and Standard Deviations for Pretraining and Posttraining Self-Ratings and Adjusted Observation Scores.

	Self-Ratings			Adjusted Observation Scores		
	Pretraining $\bar{X}$ (SD)	Posttraining $\bar{X}$ (SD)	Change	Pretraining $\bar{X}$ (SD)	Posttraining $\bar{X}$ (SD)	Change
Organization/Management						
Total (N = 40)	2.3750 (.7048)	2.8750 (.6480)	+ .5000 **	- 2.1500 (8.4476)	2.1750 (7.0707)	+ 4.3250
Elementary (N = 13)	2.0769 (.8623)	2.6154 (.7679)	+ .5385	- .1538 (8.2041)	5.3846 (3.5246)	+ 5.5385
Junior High/Middle (N = 12)	2.6667 (.4924)	3.1667 (.3892)	+ .5000	- .8333 (6.3509)	1.5833 (6.5291)	+ 2.4167
Secondary (N = 15)	2.4000 (.6325)	2.8667 (.6399)	+ .4667	- 4.9333 (9.7941)	- .1333 (8.9192)	+ 4.8000
Discipline						
Total (N = 40)	2.4750 (.9334)	2.8750 (.7906)	+ .4000	.0000 (2.9352)	.9750 (1.7757)	+ .9750
Elementary (N = 13)	2.3077 (.8549)	2.6923 (1.0316)	+ .3846	- .7692 (4.0446)	.6923 (2.1364)	+ 1.4615
Junior High/Middle (N = 12)	3.0000 (.7385)	3.2500 (.6216)	+ .2500	- .6667 (2.4985)	.5000 (2.0226)	+ 1.1667
Secondary (N = 15)	2.2000 (1.0142)	2.7333 (.5936)	+ .5333	1.2000 (1.6125)	1.6000 (.9856)	+ .4000

TABLE 1. Means and Standard Deviations for Pretraining and Posttraining Self-Ratings and Adjusted Observation Scores (continued).

	Self-Ratings			Adjusted Observation Scores		
	Pretraining $\bar{X}$ (SD)	Posttraining $\bar{X}$ (SD)	Change	Pretraining $\bar{X}$ (SD)	Posttraining $\bar{X}$ (SD)	Change
Interactive Instruction						
Total (N = 40)	2.7250 (1.3202)	3.1250 (.9111)	+ .4000	- 1.5500 (9.9870)	4.6750 (7.5018)	+ 6.2250 **
Elementary (N = 13)	2.1538 (.8987)	2.6154 (1.1929)	+ .4615	- .3077 (11.3606)	7.6923 (5.1216)	+ 8.0000
Junior High/Middle (N = 12)	2.5833 (.7930)	3.5000 (.5222)	+ .9167 **	- .9167 (8.0731)	4.3333 (8.2719)	+ 5.2500
Secondary (N = 15)	3.3333 (1.7182)	3.2667 (.7037)	- .0667	- 3.1333 (10.5618)	2.3333 (8.1299)	+ 5.4667
Monitoring Seatwork						
Total (N = 40)	2.7000 (.7860)	2.9000 (1.0080)	+ .5263 *	- 8.5500 (9.7690)	- 9.3000 (13.0839)	- .7500
Elementary (N = 13)	1.7500 (.6220)	2.1667 (1.1930)	+ .4167	- 14.6923 (12.9640)	- 17.3077 (17.9786)	- 2.6154
Junior High/Middle (N = 12)	2.6667 (.6513)	3.1667 (.7177)	+ .5000	- 6.0000 (6.2377)	- 2.9167 (3.5022)	+ 3.0833
Secondary (N = 15)	2.6429 (.7450)	3.2857 (.7260)	+ .6429	- 5.2667 (6.2503)	- 7.4667 (9.5981)	- 2.2000

* $p < .05$
** $p < .025$

TABLE 2. Analyses of Variance for Repeated Measures Summary for Observed Management and Organization

Sources of Variation	Pretraining df	Pretraining MS	Pretraining F	Pretraining p	Posttraining df	Posttraining MS	Posttraining F	Posttraining p
Total								
Pretraining Self-Rating	3	50.86	.696	.561				
Residual	36	73.07						
Posttraining Self-Rating					3	22.81	.436	.728
Residual					36	52.26		
Elementary								
Pretraining Self-Rating	3	102.09	1.832	.211				
Residual	9	55.71						
Posttraining Self-Rating					3	7.16	.505	.689
Residual					9	14.18		
Junior High/Middle								
Pretraining Self-Rating	1	12.04	.279	.609				
Residual	10	43.16						
Posttraining Self-Rating					1	190.82	6.861	.026
Residual					10	27.81		
Secondary								
Pretraining Self-Rating	2	55.04	.536	.599				
Residual	12	102.74						
Posttraining Self-Rating					2	7.49	.082	.922
Residual					12	91.56		

TABLE 3. Analyses of Variance for Repeated Measures Summary for Observed Student Discipline

	Pretraining				Posttraining			
Sources of Variation	df	MS	F	p	df	MS	F	p
Total								
Pretraining Self-Rating	3	4.93	.552	.650				
Residual	36	8.92						
Posttraining Self-Rating					3	2.17	.672	.575
Residual					36	3.24		
Elementary								
Pretraining Self-Rating	3	20.19	1.336	.323				
Residual	9	15.09						
Posttraining Self-Rating					3	2.30	.433	.735
Residual					9	5.32		
Junior High/Middle								
Pretraining Self-Rating	2	1.33	.182	.837				
Residual	9	7.33						
Posttraining Self-Rating					2	7.27	2.147	.173
Residual					9	3.39		
Secondary								
Pretraining Self-Rating	3	3.51	1.493	.271				
Residual	11	2.35						
Posttraining Self-Rating					2	.20	.182	.836
Residual					12	1.10		

TABLE 4. Analyses of Variance for Repeated Measures Summary for Observed Interactive Instruction

Sources of Variation	Pretraining df	Pretraining MS	Pretraining F	Pretraining p	Posttraining df	Posttraining MS	Posttraining F	Posttraining p
Total								
Pretraining Self-Rating	3	54.46	.518	.673				
Residual	35	105.21						
Posttraining Self-Rating					3	205.99	4.703	.007
Residual					36	43.80		
Elementary								
Pretraining Self-Rating	3	99.15	.713	.568				
Residual	9	139.04						
Posttraining Self-Rating					3	14.90	.496	.694
Residual					9	30.01		
Junior High/Middle								
Pretraining Self-Rating	2	138.45	2.832	.111				
Residual	9	48.89						
Posttraining Self-Rating					1	261.33	5.319	.044
Residual					10	49.13		
Secondary								
Pretraining Self-Rating	2	58.00	.464	.640				
Residual	11	124.99						
Posttraining Self-Rating					2	127.54	2.283	.144
Residual					12	55.86		

TABLE 5. Analyses of Variance for Repeated Measures Summary for Observed Monitoring of Seatwork

	Pretraining				Posttraining			
Sources of Variation	df	MS	F	p	df	MS	F	p
Total								
Pretraining Self-Rating	3	451.33	6.735	.001				
Residual	34	67.01						
Posttraining Self-Rating					3	472.65	3.236	.033
Residual					36	146.07		
Elementary								
Pretraining Self-Rating	2	415.08	3.921	.060				
Residual	9	105.861						
Posttraining Self-Rating					3	191.84	.523	.677
Residual					9	367.03		
Junior High/Middle								
Pretraining Self-Rating	2	14.97	.339	.722				
Residual	9	44.23						
Posttraining Self-Rating					2	7.71	..581	.579
Residual					9	13.28		
Secondary								
Pretraining Self-Rating	3	75.66	2.365	.132				
Residual	10	31.99						
Posttraining Self-Rating					2	4.76	.045	.957
Residual					12	106.69		